

Supplementary Information

Experiment	Subject group	# Blocks during 2 training phases		Delay between 2 training phases	Learning of 1st and 2 trained orientations	
		1st training	2nd training		1st trained	2nd trained
Preliminary experiment (Behavior only, Supplementary Fig. 2)	4-block training	4	N/A	N/A	No	N/A
	7-block training	7	N/A	N/A	Yes	N/A
	8-block training	8	N/A	N/A	Yes	N/A
	9-block training	9	N/A	N/A	Yes	N/A
	16-block training	16	N/A	N/A	Yes	N/A
Experiment 1 (Behavior only, Fig. 1)	No-overlearning	8	8	30 min	No	Yes
	Overlearning	16	8	30 min	Yes	No
Experiment 2 (Behavior only, Fig. 2)	No-overlearning	8	8	3.5 h	Yes	Yes
	Overlearning	16	8	3.5 h	Yes	Yes
Experiment 3 (Behavior + MRS, Fig. 3)	No-overlearning	8	N/A	N/A	Yes	N/A
	Overlearning	16	N/A	N/A	Yes	N/A
Experiment 4 (Behavior + MRS, Supplementary Fig. 9)	Overlearning	16	8	1h	Yes	No
Control Experiment 1 (Behavior only, Supplementary Fig. 3)	Overlearning	16	16	30 min	Yes	No
Control Experiment 2 (Behavior only, Supplementary Fig. 3)	No-overlearning	8	8	50 min	No	Yes
Control Experiment 3 (Behavior + MRS, Supplementary Fig. 8)	Interleaved	8 (Trained orientation 1)	8 (Trained orientation 2)	N/A	No (Trained orientation 1)	No (Trained orientation 2)
Control Experiment 4 (MRS only, Online Methods)	No training	N/A	N/A	N/A	N/A	N/A

Supplementary Table 1 Summary of the experiments and behavioral results.

Subject group	Post MRS 1 (30 min after training)		Post-MRS 2 (3.5 h after training)	
	Mean overlap (%)	s.e.m (%)	Mean overlap (%)	s.e.m (%)
No-overlearning	92.8	0.7	93.4	1.0
Overlearning	92.2	0.8	91.7	1.6

Supplementary Table 2 Mean (\pm s.e.m.) percentages of voxel overlaps between the pre-MRS and post-MRS sessions for the no-overlearning and overlearning groups in Experiment 3. The overlap percentages were not significantly different across times or groups. Results of a two-way mixed-model ANOVA on overlap percentage with main factors being time (post-MRS 1 vs. post-MRS 2 stages) and group (no-overlearning vs. overlearning groups) showed no significant main effect of either time ($F_{1,22}=0.001$, $P=0.970$) or group ($F_{1,22}=1.112$, $P=0.303$) as well as no significant interaction between the 2 factors ($F_{1,22}=0.250$, $P=0.622$).